

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/532,441
Source: PAT/10
Date Processed by STIC: 5/3/05

ENTERED



PCT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/532,441

DATE: 05/03/2005

TIME: 11:48:30

Input Set : A:\UPFT0001-100.txt
 Output Set: N:\CRF4\05032005\J532441.raw

3 <110> APPLICANT: The Trustees of the University of Pennsylvania
 5 <120> TITLE OF INVENTION: Identification of Antimycobacterial targets and the
 Inhibition
 6 Thereof as a Treatment for Infectious Diseases
 8 <130> FILE REFERENCE: UPFT0001-100
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/532,441
 C--> 10 <141> CURRENT FILING DATE: 2005-04-22
 10 <150> PRIOR APPLICATION NUMBER: PCT/US03/33524
 11 <151> PRIOR FILING DATE: 2003-10-22
 13 <150> PRIOR APPLICATION NUMBER: US 60/420,131
 14 <151> PRIOR FILING DATE: 2002-10-22
 16 <160> NUMBER OF SEQ ID NOS: 12
 18 <170> SOFTWARE: PatentIn version 3.2
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 463
 22 <212> TYPE: PRT
 23 <213> ORGANISM: M. tuberculosis H37Rv
 25 <400> SEQUENCE: 1
 27 Met Ser Pro Gln Gln Glu Pro Thr Ala Gln Pro Pro Arg Arg His Arg
 28 1 5 10 15
 31 Val Val Ile Ile Gly Ser Gly Phe Gly Gly Leu Asn Ala Ala Lys Lys
 32 20 25 30
 35 Leu Lys Arg Ala Asp Val Asp Ile Lys Leu Ile Ala Arg Thr Thr His
 36 35 40 45
 39 His Leu Phe Gln Pro Leu Leu Tyr Gln Val Ala Thr Gly Ile Ile Ser
 40 50 55 60
 43 Glu Gly Glu Ile Ala Pro Pro Thr Arg Val Val Leu Arg Lys Gln Arg
 44 65 70 75 80
 47 Asn Val Gln Val Leu Leu Gly Asn Val Thr His Ile Asp Leu Ala Gly
 48 85 90 95
 51 Gln Cys Val Val Ser Glu Leu Leu Gly His Thr Tyr Gln Thr Pro Tyr
 52 100 105 110
 55 Asp Ser Leu Ile Val Ala Ala Gly Ala Gly Gln Ser Tyr Phe Gly Asn
 56 115 120 125
 59 Asp His Phe Ala Glu Phe Ala Pro Gly Met Lys Ser Ile Asp Asp Ala
 60 130 135 140
 63 Leu Glu Leu Arg Gly Arg Ile Leu Ser Ala Phe Glu Gln Ala Glu Arg
 64 145 150 155 160
 67 Ser Ser Asp Pro Glu Arg Arg Ala Lys Leu Leu Thr Phe Thr Val Val
 68 165 170 175
 71 Gly Ala Gly Pro Thr Gly Val Glu Met Ala Gly Gln Ile Ala Glu Leu
 72 180 185 190
 75 Ala Glu His Thr Leu Lys Gly Ala Phe Arg His Ile Asp Ser Thr Lys
 76 195 200 205

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/532,441

DATE: 05/03/2005

TIME: 11:48:30

Input Set : A:\UPFT0001-100.txt

Output Set: N:\CRF4\05032005\J532441.raw

79 Ala Arg Val Ile Leu Leu Asp Ala Ala Pro Ala Val Leu Pro Pro Met
 80 210 215 220
 83 Gly Ala Lys Leu Gly Gln Arg Ala Ala Ala Arg Leu Gln Lys Leu Gly
 84 225 230 235 240
 87 Val Glu Ile Gln Leu Gly Ala Met Val Thr Asp Val Asp Arg Asn Gly
 88 245 250 255
 91 Ile Thr Val Lys Asp Ser Asp Gly Thr Val Arg Arg Ile Glu Ser Ala
 92 260 265 270
 95 Cys Lys Val Trp Ser Ala Gly Val Ser Ala Ser Arg Leu Gly Arg Asp
 96 275 280 285
 99 Leu Ala Glu Gln Ser Arg Val Glu Leu Asp Arg Ala Gly Arg Val Gln
 100 290 295 300
 103 Val Leu Pro Asp Leu Ser Ile Pro Gly Tyr Pro Asn Val Phe Val Val
 104 305 310 315 320
 107 Gly Asp Met Ala Ala Val Glu Gly Val Pro Gly Val Ala Gln Gly Ala
 108 325 330 335
 111 Ile Gln Gly Ala Lys Tyr Val Ala Ser Thr Ile Lys Ala Glu Leu Ala
 112 340 345 350
 115 Gly Ala Asn Pro Ala Glu Arg Glu Pro Phe Gln Tyr Phe Asp Lys Gly
 116 355 360 365
 119 Ser Met Ala Thr Val Ser Arg Phe Ser Ala Val Ala Lys Ile Gly Pro
 120 370 375 380
 123 Val Glu Phe Ser Gly Phe Ile Ala Trp Leu Ile Trp Leu Val Leu His
 124 385 390 395 400
 127 Leu Ala Tyr Leu Ile Gly Phe Lys Thr Lys Ile Thr Thr Leu Leu Ser
 128 405 410 415
 131 Trp Thr Val Thr Phe Leu Ser Thr Arg Arg Gly Gln Leu Thr Ile Thr
 132 420 425 430
 135 Asp Gln Gln Ala Phe Ala Arg Thr Arg Leu Glu Gln Leu Ala Glu Leu
 136 435 440 445
 139 Ala Ala Glu Ala Gln Gly Ser Ala Ala Ser Ala Lys Val Ala Ser
 140 450 455 460
 143 <210> SEQ ID NO: 2
 144 <211> LENGTH: 1392
 145 <212> TYPE: DNA
 146 <213> ORGANISM: M. tuberculosis H37Rv
 148 <400> SEQUENCE: 2
 149 atgagtcccc agcaagaacc cacagcgaa ccacccgtt ggcattcgat tttgtatcatc 60
 151 ggatctgggt tcggccggct aaacgcggca aagaagctca agcggccgca cggtgacatc 120
 153 aagctgatcg cgcgcaccac ccatcacctt ttccagccgc tgctgtacca agtggccacc 180
 155 ggatttatct ccgaggaga atcgctccg ccgaccgggg tcgtgtcgta taagcagcgc 240
 157 aatgtccagg tactttggg caacgtcacc cacatcgacc tggccggca gtgcgtcgtc 300
 159 tcggaattgc tcggcacac ctaccaaacc ccctacgaca gcctgtatgt cgccgcgggt 360
 161 gctggccagt cttatttcgg caacgaccat ttccggaaat tcgcaccgg catgaagtcc 420
 163 atcgacgacg cgttggagtt gcgtggccgc atattgagcg ctttcgagca agccgaacgg 480
 165 tccagcgatc cggAACGGCG ggccaagcta ctgacattca ccgttgcgg ggctggcccc 540
 167 accgggttgtt aaatggccgg acagatcgcc gagctggccg agcacacgtt gaaggggcgca 600
 169 ttccggcaca tcgactcgac caaggcgccg gtgattctgc ttgacgcgc cccggcggtg 660
 171 ctgccaccga tggcgcaaa gctcggtcag cggcggtcgt cccgggttgca gaagctggc 720

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/532,441

DATE: 05/03/2005
TIME: 11:48:30

Input Set : A:\UPFT0001-100.txt
Output Set: N:\CRF4\05032005\J532441.raw

173	gtggaaatcc	agctgggtgc	gatggtcacc	gacgtcgacc	gcaacggcat	caccgtcaag	780										
175	gactccgacg	gcaccgtccg	gcgcatcgag	tcggcctgca	aggctctggc	ggccggggtt	840										
177	tccggccagtc	ggttgggcag	ggaccttgcc	gagcaatcac	gggtttagct	cgaccggggcc	900										
179	ggccgggtcc	aagtgtgcc	cgacctgtcc	attcccggtt	acccgaacgt	gttcgtggtg	960										
181	ggcgatatgg	ccgctgtgga	gggtgtgcg	ggtgtggcgc	agggcgccat	ccagggggcg	1020										
183	aaatacgtcg	ccagcacgt	caaggccgaa	ctggccggcg	ccaacccggc	ggagcgtgag	1080										
185	ccattccagt	acttcgacaa	gggatcgatg	gccacggttt	cgaggtttc	ggccgtggcc	1140										
187	aagatcggtc	ccgtttagtt	cagcggcttt	atcgccctggc	tgatattggct	ggtgcgtcac	1200										
189	ctggcgtacc	tgatcggtt	caagaccaaag	atcaccactc	tgctgtcgtg	gacggtgact	1260										
191	ttcctcagta	ctcgccgccc	ccagctgacc	atcaccgacc	agcaggcatt	tgcgcgaacg	1320										
193	cggctcgaac	agctggccga	gctggccgccc	gaggcgcagg	gctcagcggc	aagcgctaag	1380										
195	gtggccagct	ag					1392										
198	<210>	SEQ ID NO:	3														
199	<211>	LENGTH:	470														
200	<212>	TYPE:	PRT														
201	<213>	ORGANISM:	M. tuberculosis H37Rv														
203	<400>	SEQUENCE:	3														
205	Met	Thr	Leu	Ser	Ser	Gly	Glu	Pro	Ser	Ala	Val	Gly	Gly	Arg	His	Arg	
206	1					5					10				15		
209	Val	Val	Ile	Ile	Gly	Ser	Gly	Phe	Gly	Gly	Leu	Asn	Ala	Ala	Lys	Ala	
210								20			25				30		
213	Leu	Lys	Arg	Ala	Asp	Val	Asp	Ile	Thr	Leu	Ile	Ser	Lys	Thr	Thr	Thr	
214								35			40				45		
217	His	Leu	Phe	Gln	Pro	Leu	Leu	Tyr	Gln	Val	Ala	Thr	Gly	Ile	Leu	Ser	
218								50			55				60		
221	Glu	Gly	Asp	Ile	Ala	Pro	Thr	Thr	Arg	Leu	Ile	Leu	Arg	Arg	Gln	Lys	
222	65							65			70				75		80
225	Asn	Val	Arg	Val	Leu	Leu	Gly	Glu	Val	Asn	Ala	Ile	Asp	Leu	Lys	Ala	
226								85			90				95		
229	Gln	Thr	Val	Thr	Ser	Lys	Leu	Met	Asp	Met	Thr	Thr	Val	Thr	Pro	Tyr	
230								100			105				110		
233	Asp	Ser	Leu	Ile	Val	Ala	Ala	Gly	Ala	Gln	Gln	Ser	Tyr	Phe	Gly	Asn	
234								115			120				125		
237	Asp	Glu	Phe	Ala	Thr	Phe	Ala	Pro	Gly	Met	Lys	Thr	Ile	Asp	Asp	Ala	
238								130			135				140		
241	Leu	Glu	Leu	Arg	Gly	Arg	Ile	Leu	Gly	Ala	Phe	Glu	Ala	Ala	Glu	Val	
242	145							145			150				155		160
245	Ser	Thr	Asp	His	Ala	Glu	Arg	Glu	Arg	Arg	Leu	Thr	Phe	Val	Val	Val	
246								165			170				175		
249	Gly	Ala	Gly	Pro	Thr	Gly	Val	Glu	Val	Ala	Gly	Gln	Ile	Val	Glu	Leu	
250								180			185				190		
253	Ala	Glu	Arg	Thr	Leu	Ala	Gly	Ala	Phe	Arg	Thr	Ile	Thr	Pro	Ser	Glu	
254								195			200				205		
257	Cys	Arg	Val	Ile	Leu	Leu	Asp	Ala	Ala	Pro	Ala	Val	Leu	Pro	Pro	Met	
258								210			215				220		
261	Gly	Pro	Lys	Leu	Gly	Leu	Lys	Ala	Gln	Arg	Arg	Leu	Glu	Lys	Met	Asp	
262	225							225			230				235		240
265	Val	Glu	Val	Gln	Leu	Asn	Ala	Met	Val	Thr	Ala	Val	Asp	Tyr	Lys	Gly	
266								245			250				255		

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/532,441

DATE: 05/03/2005
TIME: 11:48:30

Input Set : A:\UPFT0001-100.txt
Output Set: N:\CRF4\05032005\J532441.raw

269 Ile Thr Ile Lys Glu Lys Asp Gly Gly Glu Arg Arg Ile Glu Cys Ala
270 260 265 270
273 Cys Lys Val Trp Ala Ala Gly Val Ala Ala Ser Pro Leu Gly Lys Met
274 275 280 285
277 Ile Ala Glu Gly Ser Asp Gly Thr Glu Ile Asp Arg Ala Gly Arg Val
278 290 295 300
281 Ile Val Glu Pro Asp Leu Thr Val Lys Gly His Pro Asn Val Phe Val
282 305 310 315 320
285 Val Gly Asp Leu Met Phe Val Pro Gly Val Pro Gly Val Ala Gln Gly
286 325 330 335
289 Ala Ile Gln Gly Ala Arg Tyr Ala Thr Thr Val Ile Lys His Met Val
290 340 345 350
293 Lys Gly Asn Asp Asp Pro Ala Asn Arg Lys Pro Phe His Tyr Phe Asn
294 355 360 365
297 Lys Gly Ser Met Ala Thr Ile Ser Arg His Ser Ala Val Ala Gln Val
298 370 375 380
301 Gly Lys Leu Glu Phe Ala Gly Tyr Phe Ala Trp Leu Ala Trp Leu Val
302 385 390 395 400
305 Leu His Leu Val Tyr Leu Val Gly Tyr Arg Asn Arg Ile Ala Ala Leu
306 405 410 415
309 Phe Ala Trp Gly Ile Ser Phe Met Gly Arg Ala Arg Gly Gln Met Ala
310 420 425 430
313 Ile Thr Ser Gln Met Ile Tyr Ala Arg Leu Val Met Thr Leu Met Glu
314 435 440 445
317 Gln Gln Ala Gln Gly Ala Leu Ala Ala Ala Glu Gln Ala Glu His Ala
318 450 455 460
321 Glu Gln Glu Ala Ala Gly
322 465 470
325 <210> SEQ ID NO: 4
326 <211> LENGTH: 1413
327 <212> TYPE: DNA
328 <213> ORGANISM: M. tuberculosis H37Rv
330 <400> SEQUENCE: 4
331 atgacgctct catctggta accctcgccc gtcggcgcc gccatcgct ggtcatcatc 60
333 ggtagtttat tcggcgccct gaatgcggcc aaggcgctt aacggcgcca tgcgcacatc 120
335 acgctgatct ccaagacaac gaccacccctg ttccagccgc tgctgtatca agtggccacc 180
337 gggatcttgt ccgagggcga cattgccccg accacccggc tgatcctcgcc cggcaaaag 240
339 aacgtccggg tggatgttggg cgagggtcaac gcgatcgacc tgaaagcgca gacggtcacg 300
341 tcgaaaatttga tggacatgac cacggtgacg ccgtacgaca gcctcatcgat ggcggccggc 360
343 gcacagcagt cctacttcgg caacgacgaa ttccggcacct tcgcgccccg aatgaagacc 420
345 atcgacgacg cgctggagct ggcggccgc atcctggcg ctgcgcggc cgccggaggc 480
347 agcaccgacc atgccgaacg ggagcgccgc ctgacgttcg tcgtcgatcgcc cgctggccgc 540
349 accggcgctcg aggtggctgg gcagatcgatc gagctcgccg agcgacccct ggcaggcgccg 600
351 tttaggacca tcacgccccag tgagtgcggg gtatcctgc tcgcacgcgacc cccgcggtg 660
353 ttggccggcga tgggtccaaa gctgggtctc aaggcacaac ggcggcttga aaagatggac 720
355 gtcgaggttc aactcaacgc gatggtgacc gcggtcgact acaaaggcat caccatcaag 780
357 gaaaaggacg gcccggaaacg ccgcacatcgaa tgcgcgtgca aggtttgggc ggcggcgctg 840
359 gcccggcagcc cgctggccaa gatgatcgcc gagggatccg acggAACCGA aatcgaccgg 900
361 gcccggaaagggg tgatcgttggaa acccgatctc accgtcaagg gacatccgaa cgtcttcgtt 960

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/532,441

DATE: 05/03/2005
TIME: 11:48:30

Input Set : A:\UPFT0001-100.txt
Output Set: N:\CRF4\05032005\J532441.raw

363	gtcggcgatc	tgatttcg	ccccggcgta	cccggggtgg	ctcagggcgc	gatccagggg	1020									
365	gccccatacg	ccaccacgg	gtcaaaacac	atggtcaagg	gcaatgacga	cccagccaat	1080									
367	cgcggccgt	tccattactt	caacaaggc	agcatggcga	cgcgttcccg	ccacagcgcc	1140									
369	gtcggcgagg	tggcaagct	ggagtttgc	gggtacttcg	cctggctggc	gtggctggtg	1200									
371	ctgcacctgg	tctacctgg	cggctatcg	aaccgcac	cagccctgtt	cgcttggggg	1260									
373	atcccttca	tggccgcgc	ccggggccag	atggccatca	ccagccagat	gatctacgcc	1320									
375	agtttagtga	tgaccttgat	gaaacagcag	gcacaaggag	cgctggcagc	cgccgaacag	1380									
377	gcccggc	ccgagcaaga	ggcagcgggt	tag			1413									
380	<210>	SEQ ID NO:	5													
381	<211>	LENGTH:	219													
382	<212>	TYPE:	PRT													
383	<213>	ORGANISM:	M. tuberculosis	CDC1551												
385	<400>	SEQUENCE:	5													
387	Met	Phe	Asp	Gly	Val	Ala	Arg	Lys	Tyr	Asp	Leu	Thr	Asn	Thr	Val	Leu
388	1				5					10						15
391	Ser	Leu	Gly	Gln	Asp	Arg	Tyr	Trp	Arg	Arg	Ala	Thr	Arg	Ser	Ala	Leu
392						20				25						30
395	Arg	Ile	Gly	Pro	Gly	Gln	Lys	Val	Leu	Asp	Leu	Ala	Ala	Gly	Thr	Ala
396						35				40						45
399	Val	Ser	Thr	Val	Glu	Leu	Thr	Lys	Ser	Gly	Ala	Trp	Cys	Val	Ala	Ala
400						50				55						60
403	Asp	Phe	Ser	Val	Gly	Met	Leu	Ala	Ala	Gly	Ala	Ala	Arg	Lys	Val	Pro
404						65				70						80
407	Lys	Val	Ala	Gly	Asp	Ala	Thr	Arg	Leu	Pro	Phe	Gly	Asp	Asp	Val	Phe
408						85				90						95
411	Asp	Ala	Val	Thr	Ile	Ser	Phe	Gly	Leu	Arg	Asn	Val	Ala	Asn	Gln	Gln
412						100				105						110
415	Ala	Ala	Leu	Arg	Glu	Met	Ala	Arg	Val	Thr	Arg	Pro	Gly	Gly	Arg	Leu
416						115				120						125
419	Leu	Val	Cys	Glu	Phe	Ser	Thr	Pro	Thr	Asn	Ala	Leu	Phe	Ala	Thr	Ala
420						130				135						140
423	Tyr	Lys	Glu	Tyr	Leu	Met	Arg	Ala	Leu	Pro	Arg	Val	Ala	Arg	Ala	Val
424						145				150						160
427	Ser	Ser	Asn	Pro	Glu	Ala	Tyr	Glu	Tyr	Leu	Ala	Glu	Ser	Ile	Arg	Ala
428						165				170						175
431	Trp	Pro	Asp	Gln	Ala	Val	Leu	Ala	His	Gln	Ile	Ser	Arg	Ala	Gly	Trp
432						180				185						190
435	Ser	Gly	Val	Arg	Trp	Arg	Asn	Leu	Thr	Gly	Gly	Ile	Val	Ala	Leu	His
436						195				200						205
439	Ala	Gly	Tyr	Lys	Pro	Gly	Lys	Gln	Thr	Pro	Gln					
440						210				215						
443	<210>	SEQ ID NO:	6													
444	<211>	LENGTH:	660													
445	<212>	TYPE:	DNA													
446	<213>	ORGANISM:	M. tuberculosis	CDC1551												
448	<400>	SEQUENCE:	6													
449	atgttcgatg	gcgtcgcccg	caagtatgac	ctgaccaata	ccgtgttgc	cctggccag										60
451	gaccggatt	ggcgccgagc	cactcggtcg	gctgtcgga	tccggcccg	ccaaaaggtc										120
453	ctggacctgg	ccgcgggcac	cgcgtgtcc	accgtagac	tcaccaaatac	ggcgcggtgg										180

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/532,441

DATE: 05/03/2005

TIME: 11:48:31

Input Set : A:\UPFT0001-100.txt

Output Set: N:\CRF4\05032005\J532441.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date